£.3

ra Ca

fe i fe

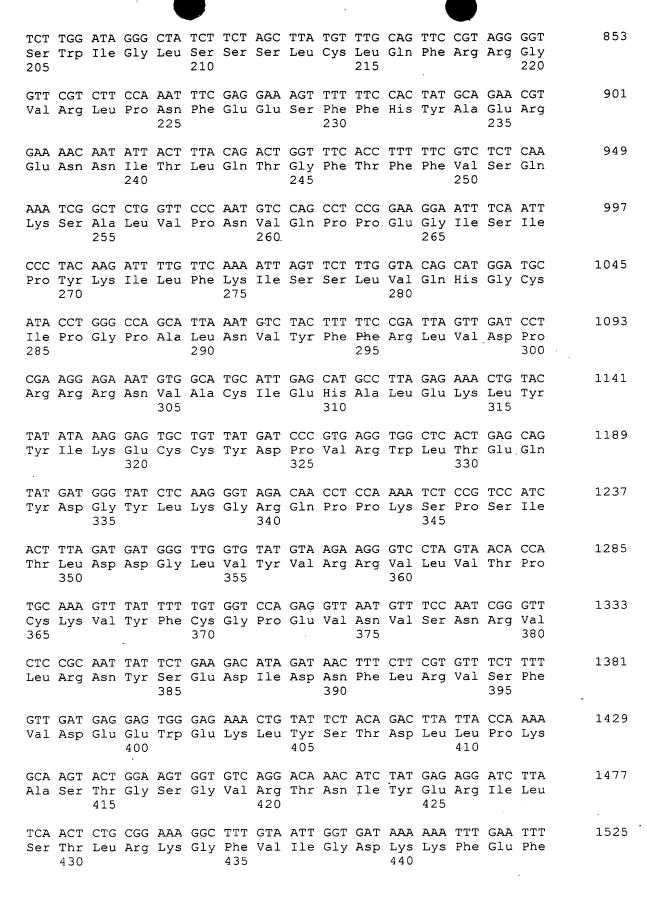
SEQUENCE LISTING

(1) GENERAL INFORMATION:

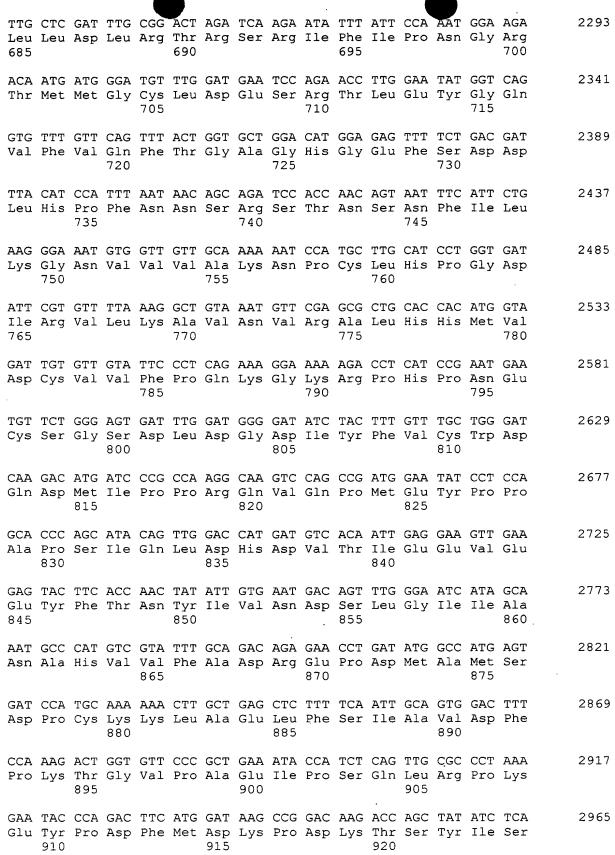
- (i) APPLICANT: Wassenegger, Michael Riedel, Leonhard Schiebel, Winfried Sanger, Heinz
- (ii) TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING POLYPEPTIDES HAVING THE ENZYMATIC ACTIVITY OF AN RNA-DIRECTED RNA POLYMERASE (RdRP)
- (iii) NUMBER OF SEQUENCES: 13
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: FISH & NEAVE
 - (B) STREET: 1251 Avenue of the Americas
 - (C) CITY: New York
 - (D) STATE: New York
 - (E) COUNTRY: USA
 - (F) ZIP: 10020
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/811,583
 - (B) FILING DATE: 05-MAR-1997
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Haley, James F.
 - (B) REGISTRATION NUMBER: 27,794
 - (C) REFERENCE/DOCKET NUMBER: MPG-1
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 212-596-9000
 - (B) TELEFAX: 212-596-9090
- (2) INFORMATION FOR SEQ ID NO:1:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3731 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: cDNA
 - (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Tomato
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 194..3535

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAAATATTCT I	TACTTACTT C	ACCAGGGAT TO	SACTCATCA CTO	CCCCTCAA GTCT	TTGTGT 60
GTTGTGATAA T	TAAATTTGGT T	GTGCTTCAG TI	TTCAGTCAC TAC	CTGCTGGG TAGT	TTTTAT 120
TTTGCATAAC T	TTCAGGGGGT A	TTCCAGTTG GT	GTTAGCAT TTO	GAAAGTCG AACT	GCACTT 180
GGAATTTGGC I				GGA TTC CCT T Gly Phe Pro T	
				G AAA TAT ACA u Lys Tyr Thr 25	
				C AAA GGA GGA r Lys Gly Gly 0	
				A AGT GCT GAC e Ser Ala Asp	
				C TCT TCT TAT y Ser Ser Tyr 75	
			e Val Gln Leu	G CGG GCA TAT u Arg Ala Tyr 90	
				T CAG ATA TCA s Gln Ile Ser 105	
				A ATT CAA TTT r Ile Gln Phe O	
				T GGT TCA GCT r Gly Ser Ala	
				G GTT GTG CTC n Val Val Leu 155	
			e Leu Leu Ile	A CAG TTA TTT e Gln Leu Phe 170	
				T TAT AGC TTC s Tyr Ser Phe 185	
				A GAT TTC CCT r Asp Phe Pro 0	



	GCA Ala									1573
	TCA Ser									1621
	TTT Phe									1669
	TCT Ser									1717
	GAA Glu 510									1765
	GAT Asp									1813
	AAA Lys									1861
	GGA Gly									1909
	TCT Ser									1957
	GAT Asp 590									2005
_	CTG Leu	_								2053
	CAG Gln									2101
	TCT									2149
	ACT Thr									2197
	GAG Glu 670									2245



GAA AGA GTT ATT GGA AAG CTT TTC AGG AAA GTG AAG GAC AAA GCA CCT Glu Arg Val Ile Gly Lys Leu Phe Arg Lys Val Lys Asp Lys Ala Pro 925 930 935 940	3013							
CAG GCT AGC TCT ATC GCG ACC TTC ACA AGA GAT GTT GCA AGG AGA TCA Gln Ala Ser Ser Ile Ala Thr Phe Thr Arg Asp Val Ala Arg Arg Ser 945 950 955	3061							
TAT GAT GCT GAT ATG GAA GTT GAT GGA TTT GAA GAT TAC ATT GAC GAA Tyr Asp Ala Asp Met Glu Val Asp Gly Phe Glu Asp Tyr Ile Asp Glu 960 965 970	3109							
GCT TTT GAC TAC AAA ACT GAA TAT GAC AAC AAG CTG GGT AAT TTA ATG Ala Phe Asp Tyr Lys Thr Glu Tyr Asp Asn Lys Leu Gly Asn Leu Met 975 980 985	3157							
GAC TAC TAT GGC ATA AAA ACA GAG GCT GAA ATA CTT AGT GGT GGC ATT ASP Tyr Tyr Gly Ile Lys Thr Glu Ala Glu Ile Leu Ser Gly Gly Ile 990 995 1000	3205							
ATG AAG GCA TCA AAA ACT TTT GAC CGC AGA AAA GAT GCT GAG GCC ATT Met Lys Ala Ser Lys Thr Phe Asp Arg Arg Lys Asp Ala Glu Ala Ile 1005 1010 1015 1020	3253							
AGT GTT GCT GTG AGG GCC TTG AGG AAG GAG GCA AGA GCC TGG TTC AAG Ser Val Ala Val Arg Ala Leu Arg Lys Glu Ala Arg Ala Trp Phe Lys 1025 1030 1035	3301							
AGG CGT AAT GAT ATA GAT GAC ATG TTA CCA AAG GCT TCG GCT TGC TAC Arg Arg Asn Asp Ile Asp Asp Met Leu Pro Lys Ala Ser Ala Trp Tyr 1040 1045 1050	3349							
CAC GTT ACA TAT CAT CCT ACA TAT TGG GGT TGC TAC AAT CAG GGG TTG His Val Thr Tyr His Pro Thr Tyr Trp Gly Cys Tyr Asn Gln Gly Leu 1055 1060 1065	3397							
AAA AGA GCT CAT TTC ATT AGC TTT CCC TGG TGT GTT TAT GAC CAG CTA Lys Arg Ala His Phe Ile Ser Phe Pro Trp Cys Val Tyr Asp Gln Leu 1070 1075 1080	3445							
ATC CAG ATT AAG AAG GAC AAA GCA CGT AAC AGG CCA GTT CTC AAC TTG Ile Gln Ile Lys Lys Asp Lys Ala Arg Asn Arg Pro Val Leu Asn Leu 1085 1090 1095 1100	3493							
TCA TCT CTC AGG GCT CAA CTG AGT CAC AGA TTA GTG TTG AAA Ser Ser Leu Arg Ala Gln Leu Ser His Arg Leu Val Leu Lys 1105	3535							
TGAGATTCCA GTCGAGCGTT AAGCTGATAT ATATATAATG TAATAGGGTG TGATCATAAG	3595							
AAAACTGTTA TGCATTGTTG ACTACCTTTT GTCTTTAAAA CTGCATGAAG CTGCAACATA	3655							
TATGCAGTAC TCTAAGAAAC AGATGTACAG CTAAGTACTA ATATGTATGT GATTTGAGTT								
TCATCTTTCT TCTAAA								

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 1114 amino acids



(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly Lys Thr Ile Gln Val Phe Gly Phe Pro Tyr Leu Leu Ser Ala 1 5 10 15

Glu Val Val Lys Ser Phe Leu Glu Lys Tyr Thr Gly Tyr Gly Thr Val 20 25 30

Cys Ala Leu Glu Val Lys Gln Ser Lys Gly Gly Ser Arg Ala Phe Ala 35 40 45

Lys Val Gln Phe Ala Asp Asn Ile Ser Ala Asp Lys Ile Ile Thr Leu 50 55 60

Ala Asn Asn Arg Leu Tyr Phe Gly Ser Ser Tyr Leu Lys Ala Trp Glu 65 70 75 80

Met Lys Thr Asp Ile Val Gln Leu Arg Ala Tyr Val Asp Gln Met Asp 85 90 95

Gly Ile Thr Leu Asn Phe Gly Cys Gln Ile Ser Asp Asp Lys Phe Ala 100 105 110

Val Leu Gly Ser Thr Glu Val Ser Ile Gln Phe Gly Ile Gly Leu Lys 115 120 125

Lys Phe Phe Phe Leu Ser Ser Gly Ser Ala Asp Tyr Lys Leu Gln 130 135 140

Leu Ser Tyr Glu Asn Ile Trp Gln Val Val Leu His Arg Pro Tyr Gly 145 150 155 160

Gln Asn Ala Gln Phe Leu Leu Ile Gln Leu Phe Gly Ala Pro Arg Ile 165 170 175

Tyr Lys Arg Leu Glu Asn Ser Cys Tyr Ser Phe Phe Lys Glu Thr Pro 180 185 190

Asp Asp Gln Trp Val Arg Thr Thr Asp Phe Pro Pro Ser Trp Ile Gly 195 200 205

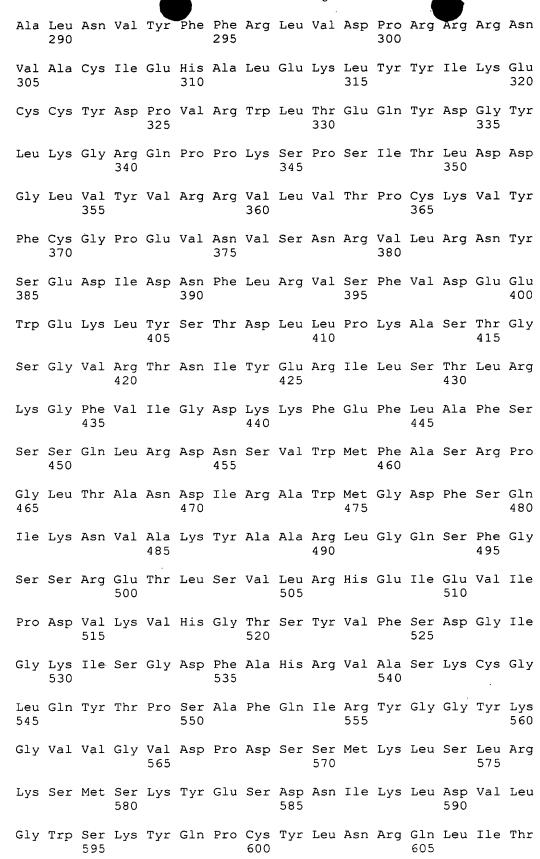
Leu Ser Ser Ser Leu Cys Leu Gln Phe Arg Arg Gly Val Arg Leu Pro 210 215 220

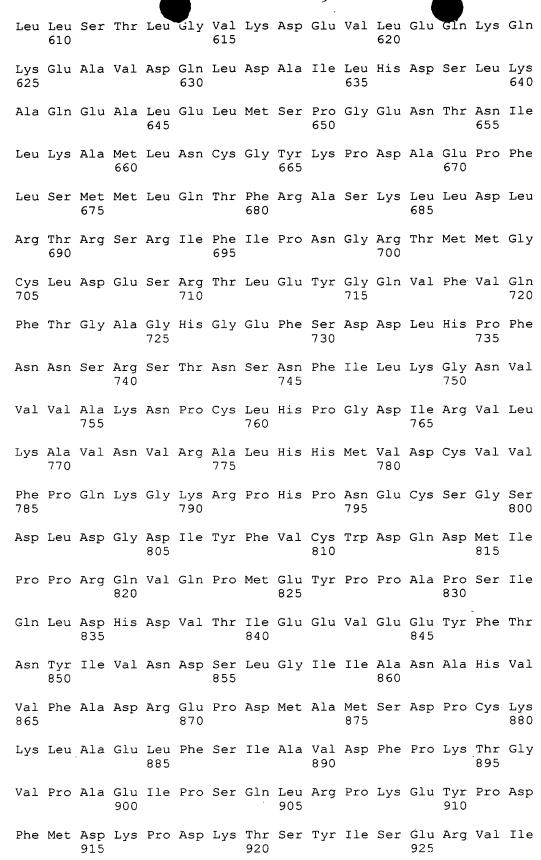
Asn Phe Glu Glu Ser Phe Phe His Tyr Ala Glu Arg Glu Asn Asn Ile 225 230 235 240

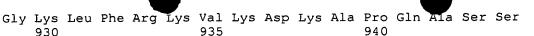
Thr Leu Gln Thr Gly Phe Thr Phe Phe Val Ser Gln Lys Ser Ala Leu 245 250 255

Val Pro Asn Val Gln Pro Pro Glu Gly Ile Ser Ile Pro Tyr Lys Ile 260 265 270

Leu Phe Lys Ile Ser Ser Leu Val Gln His Gly Cys Ile Pro Gly Pro 275 280 285







Ile Ala Thr Phe Thr Arg Asp Val Ala Arg Arg Ser Tyr Asp Ala Asp 945 950 955 960

Met Glu Val Asp Gly Phe Glu Asp Tyr Ile Asp Glu Ala Phe Asp Tyr 965 970 975

Lys Thr Glu Tyr Asp Asn Lys Leu Gly Asn Leu Met Asp Tyr Tyr Gly 980 985 990

Ile Lys Thr Glu Ala Glu Ile Leu Ser Gly Gly Ile Met Lys Ala Ser 995 1000 1005

Lys Thr Phe Asp Arg Lys Asp Ala Glu Ala Ile Ser Val Ala Val 1010 1015 1020

Arg Ala Leu Arg Lys Glu Ala Arg Ala Trp Phe Lys Arg Arg Asn Asp 1025 1030 1035 1040

Ile Asp Asp Met Leu Pro Lys Ala Ser Ala Trp Tyr His Val Thr Tyr 1045 1050 1055

His Pro Thr Tyr Trp Gly Cys Tyr Asn Gln Gly Leu Lys Arg Ala His 1060 1065 1070

Phe Ile Ser Phe Pro Trp Cys Val Tyr Asp Gln Leu Ile Gln Ile Lys 1075 1080 1085

Lys Asp Lys Ala Arg Asn Arg Pro Val Leu Asn Leu Ser Ser Leu Arg 1090 1095 1100

Ala Gln Leu Ser His Arg Leu Val Leu Lys 1105 1110

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 218 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Arg Thr Met Met Gly Cys Leu Asp Glu Ser Arg Thr Leu Glu Tyr Gly
1 10 15

Gln Val Phe Val Gln Phe Thr Gly Ala Gly His Gly Glu Phe Ser Asp 20 25 30

Asp Leu His Pro Phe Asn Asn Ser Arg Ser Thr Asn Ser Asn Phe Ile $35 \hspace{1cm} 40 \hspace{1cm} 45$

Leu Lys Gly Asn Val Val Val Ala Lys Asn Pro Cys Leu His Pro Gly 50 60



Asp Ile Arg Val Leu Lys Ala Val Asn Val Arg Ala Leu His His Met 65 70 75 80

Val Asp Cys Val Val Phe Pro Gln Lys Gly Lys Arg Pro His Pro Asn 85 90 95

Glu Cys Ser Gly Ser Asp Leu Asp Gly Asp Ile Tyr Phe Val Cys Trp 100 105 110

Asp Gln Asp Met Ile Pro Pro Arg Gln Val Gln Pro Met Glu Tyr Pro 115 120 125

Pro Ala Pro Ser Ile Gln Leu Asp His Asp Val Thr Ile Glu Glu Val 130 135 140

Glu Glu Tyr Phe Thr Asn Tyr Ile Val Asn Asp Ser Leu Gly Ile Ile 145 150 155 160

Ala Asn Ala His Val Val Phe Ala Asp Arg Glu Pro Asp Met Ala Met 165 170 175

Ser Asp Pro Cys Lys Lys Leu Ala Glu Leu Phe Ser Ile Ala Val Asp 180 185 190

Phe Pro Lys Thr Gly Val Pro Ala Glu Ile Pro Ser Gln Leu Arg Pro 195 200 205

Lys Glu Tyr Pro Asp Phe Met Asp Lys Pro 210 215

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "primer"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CATAACGAAT CTGGAAAGCA GATGG

- (2) INFORMATION FOR SEQ ID NO:5:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "primer"

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:	
GATGAATCCG GATCAACACC CACAC	25
(2) INFORMATION FOR SEQ ID NO:6:	
 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 25 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
<pre>(ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: /desc = "primer"</pre>	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
GGGTGCTGGA GGATATTCCA TCGGC	25
(2) INFORMATION FOR SEQ ID NO:7:	
(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 34 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
<pre>(ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: /desc = "primer"</pre>	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:	
CTTCACCAGG GATCCACTCA TCACTCCCCT CAAG	34
(2) INFORMATION FOR SEQ ID NO:8:	
 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 34 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
<pre>(ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: /desc = "primer"</pre>	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:	
GCATAACTTC AGGGGGGATC CAGTTGGTGT TAGC	34

(2) INFORMATION FOR SEQ ID NO:9:





- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "primer"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCAGCTTCAT GCAGATCTAA AGACAAAAGG TAGTC

35

- (2) INFORMATION FOR SEQ ID NO:10:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

- (2) INFORMATION FOR SEQ ID NO:11:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: peptide
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

- (2) INFORMATION FOR SEQ ID NO:12:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear

(ii) MOLECULE TIPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ser

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 13 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Val Phe Pro Gln Lys Gly Lys Arg Pro His Asn Glu Cys 1 5 10